Journal of Neuroscience, Psychology, and Economics
2014, Vol. 7, No. 1, 15–32

1937-321X/14/$12.00 DOI: 10.1037/npe0000017

A Dynamic Model of Dissonance Reduction in a Modular Mind

Emin Karagözoğlu
Bilkent University

This article establishes an economic framework to identify the conditions under which an optimizing agent may seek (or not seek) to engage in costly dissonance reduction. We set up a dynamic model of decision-making in which the individual's mind is composed of a coordinating principal and 2 conflicting agents. We take into account the *cognitive dissonance* experienced as a result of the conflict between the agents. Each agent (or *self*) is specialized in perceiving a particular type of signal and performing a task. Dissonance levels in our model are not constant. Instead, the individual's split-selves are open to *habituation,* which can lead to a reduction in cognitive dissonance. Therefore, *motivated habituation* appears as a way to avoid dissonance. We prove the existence of an optimal strategy with a threshold structure. Our results show that the existence of intrapersonal conflict may be a long run phenomenon even in an optimizing mind.

Keywords: action-based theory, cognitive dissonance, habituation, mind hierarchy, modularity of mind, stochastic dynamic programming

In this article we seek to investigate the decision-making behavior of an individual whose choices are partially constrained by *cognitive dissonance*. Our research questions are as follows: (i) What is the behavior induced by an optimizing mind in a dynamic setting in which there are states of the world that cause cognitive dissonance and correspondingly a fall in decision-making performance? (ii) What is the evolution of cognitive dissonance on an optimal decision path? and (iii) What are the factors that affect the dissonance reduction decision of an optimizing mind?

This article was published Online First December 23, 2013.

Emin Karagözoğlu, Department of Economics, Bilkent University, Ankara, Turkey.

An earlier version of this work was circulated under the title "A Dynamic Model Decision-Making Under Cognitive Dissonance and Modularity of Mind." I thank an editor and two anonymous reviewers for constructive comments and critiques that improved the paper. I also thank Kalyan Chatterjee and Arno Riedl for valuable comments and discussions. Finally, I owe thanks to Zafer Akın, Simon Gächter, İsa Emin Hafalır, Ayşe Tever Karagözoğlu, Jeffrey Kurland, and Abdullah Yavaş for their helpful comments. Usual disclaimers apply.

Correspondence concerning this article should be addressed to Emin Karagözoğlu, Department of Economics, Bilkent University, 06800 Ankara, Turkey. E-mail: karagozoglu@bilkent.edu.tr

A motivation for our study stems from the observation that stress and tension attributable to conflicting beliefs or actions are experienced more frequently, repetitively, and over a longer horizon than one would expect (see Mahaffy, 1996; Zhou, 2000; Fischbacher and Heusi, 2008; Gino, Ayal and Ariely, 2009; Gino, Norton and Ariely, 2010; Garvey, 2012). Given that cognitive dissonance is not a pleasant state and individuals try to avoid it, the above-mentioned observation raises a question: Can long-run dissonance (or *living with dissonance* as mentioned in Mahaffy, 1996; Zhou, 2000; Garvey, 2012 and as *living in contradiction* in Margolin, 1997) be optimal? If it is, which factors have an impact on the *long run* (optimal) level of conflict an individual bears in the mind? What is the influence and the role of *habituation* on the evolution of the level of intrapersonal conflict?[1] To answer these questions, we model the mind as a decision-making unit with a modular structure, which is aware of the dynamic and flexible aspect of the (potential) conflict between its different parts and acts accordingly. In particular, it may forego higher returns in some states

[1] Habituation is defined as the decline of a conditioned response after repeated exposure to the conditioned stimulus. More discussion will follow in The Model.

of the world to be able to reduce the cognitive dissonance experienced due to the conflict that occurs in some other states of the world. The concept of *habituation,* which in the current context implies that cognitive dissonance may change as a result of being repetitively exposed to the dissonance inducing state/signal, is a corner stone of our model. The dynamic character of the model allows us to analyze the impact of relevant parameters on the evolution of intrapersonal conflict.

Our model incorporates the approach that the mind is not a *seamless* unit, but rather a *modular* one (see Fodor, 1983; Minsky, 1986; Cosmides and Tooby, 1992). It has different parts/systems in a hierarchical structure, which are specialized in certain tasks and responsive to only particular signals (see Fodor, 1983; Carruthers, 2006; Samuels, 2006). In our model, there is an organizing unit (the *principal*, resembling the *central system*, that deals with complex cognitive activities) and, for the sake of simplicity, two different units in a lower hierarchy (*agents*, resembling *input systems*, that deal with basic cognitive activities). Each agent has a comparative advantage in receiving a signal. When the signal describing the world arrives, the principal assigns it to one (or in some cases both) of the agents who then performs a task. For our purposes, this task can be interpreted as a simple binary decision. If the signal is perceived successfully, then the agent performs the task, which brings a positive return (or utility) to the individual. The reception of the signal and the execution of the task can even be considered together as one activity. It is worthwhile emphasizing that cognitive dissonance arises only in the case of certain states of the world, in which the information asymmetry between (potentially) conflicting units/cognitions vanishes, which is a difference from the standard definition of the concept.[2]

The principal's decision in each period can be reduced to choosing one of the following actions (or choices): attempting *cross-training* or acting in accordance with the agents' specializations. Cross-training is a method of costly dissonance reduction: it aims at reducing the dissonance by exposing the agents to characteristics of the world they do not "like" and/or are not specialized in perceiving. There is a current cost associated with *cross-training*, because it relies on assigning a signal to an agent that is less likely to perceive this signal. Nevertheless, cross-training may lead to lower dissonance levels in the future because it has the potential to generate *habituation.* The return from cross-training accrues in the composite state of the world, when the principal needs both agents to act together, which is attributable to the fact that cognitive dissonance is felt when these two conflicting agents have to act together (see Fodor, 1983; Minsky, 1986; Brocas and Carillo, 2008a, 2008b for information encapsulation and asymmetric information in the mind and brain). As the reader may notice, the negative effects of cognitive dissonance in our model are reflected on task performance or effectivity of actions, which is more in line with the action-based theory of cognitive dissonance (Harmon-Jones, 1999) than the standard theory of cognitive dissonance (Festinger, 1957). Acting in accordance with agents' specializations is just assigning the signal to the agent who is specialized in it (or more likely to perceive it). Acting in accordance with specializations leads to a higher expected current return, but does not reduce dissonance levels. Therefore, there is a trade-off associated with current and future returns.

The elements we explain above define a stochastic dynamic programming problem. Solving this problem, we find that a forward-looking principal uses a *threshold rule* in deciding whether to engage in costly cognitive dissonance reduction or not. An implication of this finding is that for some parameter values, it is *not* optimal to totally avoid dissonance. Therefore, our main result implies that under certain circumstances, which will be described in terms of the parameters of the model, the intrapersonal conflict may remain in an optimizing mind that is, cognitive dissonance can be a long-run (optimal) phenomenon. Individual's defining characteristics and observed behavior are affected by the comparative flexibility of his different selves. The pace of habituation, the time preference, and the frequency of the dissonance inducing state are some important factors influencing the optimal decision. In partic-

[2] I thank an anonymous reviewer for raising a question, which lead to this clarifying statement.

ular, an increase in the pace of habituation, the discount parameter (i.e., forwardlookingness), and the frequency of conflict have the same kind of impact: they all increase the expected return from attempting cross-training, ultimately reducing the long run level of dissonance.

Most of the contributions of our model lie in its simple yet rich behavioral structure. First of all, to the best of our knowledge, the model in this article is the first dynamic model of individual decision-making under cognitive dissonance. Second, it is a model that incorporates many behavioral phenomena such as habituation, modularity of and hierarchy in the mind, and cognitive dissonance to an individual decision-making problem. Finally, the current study offers *motivated habituation* as an alternative way to avoid cognitive dissonance. Some other methods that existing literature offer are moral disengagement, attitude change, motivated forgetting, selective exposure, and denial of responsibility. Among these, motivated habituation is closer to motivated forgetting or attitude change. Nevertheless, it is still different from them in that motivated habituation is an ex-ante action taken to avoid the negative effects of future, anticipated cognitive dissonance, whereas the former ones are ex-post measures that aim to avoid (current or already experienced) dissonance. It is the dynamic character of our model, which enables a method that aims at reducing future, anticipated dissonance.

Literature Review

In his seminal piece, Festinger (1957) asserts that people experience cognitive dissonance when they hold inconsistent beliefs or act contrary to their existing beliefs. Because the tension brought by this inconsistency between different cognitions is not a pleasant state, individuals want to avoid it.[3] Aronson (1988) and Beauvois and Joule (1996) claim that dissonance is mainly related to the issues of self-esteem and the desire to rationalize one's actions. In other words, the dissonance and dissonance reducing behavior generally exist when an individual's actions or beliefs are in conflict with his or her desire to be known/seen as a good or an intelligent individual. This argument is complementing Festinger's (1957) original theory of cognitive dissonance in a way that it emphasizes the main source of dissonance.

Festinger's (1957) theory of cognitive dissonance is considered as one of the most influential theories in psychology (see Jones, 1985). It led to thousands of articles in the last six decades. Naturally, we can mention here only the studies that are closely related to ours. This means we focus on the ones using mathematical techniques to model cognitive dissonance reduction. Shultz and Lepper (1996) develop a computational model of cognitive dissonance, which they call *consonance model.* They use networks to model individuals' descriptions of decision problems: node directions and weights refer to relationships between different cognitions. Using this model they simulate data and compare the simulated data to human behavior in some well-known cognitive dissonance experiments. A static version of our model with only the dissonance inducing state would be similar to a simple version of their network model. Sakai (1999) develops a multiplicative power-function model of cognitive dissonance. His model can deliver predictions about the emergence and level of cognitive dissonance. One main difference between Sakai (1999) and our model is that we do not model or explain the emergence of cognitive dissonance.

Lévy-Garboua and Blondel (2002) argue that cognitive dissonance does not presume irrational behavior if one defines the rationality notion properly under dynamic uncertainty. They argue that cognitive consistency must be the rationality notion used under dynamic uncertainty. Hence, a common, intuitive message of our article and Lévy-Garboua and Blondel (2002) is that cognitive dissonance and rationality are less incompatible than usually thought. Livnat and Pippenger (2006) set up a game theoretical model of evolution and show that the presence of conflicting agents in an optimal brain is possible, which is again in line with our main result.

[3] Festinger (1957) argues "The existence of dissonance, being psychologically uncomfortable, will motivate the person to try to reduce the dissonance and achieve consonance."

Gawronski and Kulakowski (2007) generalize the differential equations approach used to model cognitive dissonance in groups to consider asymmetric interpersonal ties. Their study follows a Heiderian approach to cognitive dissonance in social groups (see Heider, 1946). Their numerical results show that under certain circumstances, it may take arbitrarily long time to avoid intragroup dissonance. Along similar lines with Shultz and Lepper (1996) and Sakai (1999), Sakai (2013) proposes a recurrent network model of cognitive consistency to unify Heider's balance theory with Festinger's cognitive dissonance theory. One common difference between our article and most of the articles mentioned above is that we solve our model analytically whereas the others come up with computational solutions and simulations.

The optimization approach and dynamic mathematical models are more widespread in economics. On the other hand, research on cognitive dissonance in economics started more than two decades after Festinger (1957). The first article in economics that deals with possible economic consequences of cognitive dissonance is Akerlof and Dickens (1982). Their model predicts not only how given information is interpreted, but also whether that information is to be received or not according to their preferences. With the cognitive dissonance approach, they can explain the effectivity of noninformational advertising, the popularity of social security legislation, and safety legislation and the failure of people in purchasing actuarially beneficial flood and earthquake insurance. Dickens (1986), Gilad et al. (1987), Rabin (1994), Brady et al. (1995), Konow (2000), Oxoby (2003), Goldsmith et al. (2004), Balestrino and Ciardi (2008), Johansson-Stenman and Svedsäter (2008), Bendersky and Curhan (2009), Lester and Young (2009), Smith (2009), Ishida (2010), Dickinson and Oxoby (2011), and Matthey and Regner (2011) are some other articles that use cognitive dissonance to explain certain economic phenomena or behavior such as criminal behavior, sunk cost fallacy, status-seeking behavior, labor supply responses to discrimination, ultimatum game offers, negotiation behavior, principal–agent relationships, timing of marriage, pessimism, other-regarding behavior, and behavioral spillover effects.

Epstein and Kopylov (2006), to the best of our knowledge, is the only axiomatic study of cognitive dissonance in decision theory. These authors model self-justification of past decisions. Axioms are defined on preferences over ex-ante actions (modeled formally by menus). The representation of these preferences admits the interpretation that the agent adjusts beliefs after taking an action so as to be more optimistic about its possible consequences. Since the action has already been taken, no adjustment can be made on it. Therefore, the individual adjusts his or her beliefs taking actions as given. Brocas and Carillo (2008a, 2008b) and Alonso, Brocas, and Carillo (2011) model brain activity and provide neuroeconomic foundations for variety of behavior. Our article is comparable, in its nature, to these last few studies more than the others in economics mentioned above in the sense that our primary goal, as well, is modeling a behavioral phenomenon (e.g., dissonance reduction) rather than using it to explain an economic phenomena.

To summarize, the current article lies at the intersection of psychology and economics in that it aims to explain a psychological phenomenon with mathematical techniques frequently used in economics. What makes our model different from earlier studies mentioned above is its dynamic and rich structure, both of which enable us to analyze the effects of some parameters relevant for costly dissonance avoidance and offer motivated habituation as a resolution strategy.

The Model

In this section, we present our model and explain each modeling assumption by referring to relevant studies in the literature.

Modularity of Mind

Our model of decision-making in the human mind stems from ideas and findings in two literatures. Findings in neuroscience and cognitive science literature suggest that the mind is not a *seamless* unit. Rather, it is composed of *semi-autonomous* parts each of which is specialized in receiving certain signals and performing certain tasks. This implies *modularity* of the mind in cognitive activity. One of the major points Minsky's (1986) *Society of Mind* makes is this *collective but specialized* structure of human mind. Fodor (1983); Cosmides and Tooby (1987), Sperber (1994), Pinker (1997), Shettleworth (1998), and Carruthers (2006) are some other prominent proponents of the modular mind argument.

A well-known example in neuroscience is attributable to McClure et al. (2004), reporting that the limbic system responds more to immediate gratification whereas the cortical system responds to longer-term benefits. Sergent and Signoret (1992); Charman and Baron-Cohen (1995), Miller and Cohen (2001), Geary and Huffman (2002), Poldrack and Rodriguez (2004), and Bechara (2005) are some other neuroscientific studies reporting on the modularity of (or compartmentalization in) brain. Finally, Livnat and Pippenger (2006) and Bisin and Iantchev (2010) present evolutionary arguments in favor of modularity.

To summarize, the modularity of mind (or brain) is a prominent approach embraced in cognitive psychology, evolutionary psychology, and neuroscience literature.[4] In economics, the multiple-selves approach, which is in line with this phenomenon, has been used extensively in recent years. People have different traits, inner selves, or characteristics that are specialized in different signals, tasks, and responsive to different types of urges.

Hierarchical Structure

Another important finding that is relevant for our model is that there is a *hierarchy* in the mind. Some systems or parts have superior positions in this decision-making hierarchy (see Hughlings-Jackson, 1959; Duncan and Owen, 2000; Botvinick et al., 2001; Miller and Cummings, 2007; Badre and D'Esposito, 2007). Accordingly, we assume that there is a governing unit, the *principal,* and two semi-autonomous *agents* embedded in the mind. This is also similar to the *central system* versus *input systems* in Fodor (1983), *planner-doer* interpretation in Thaler and Shefrin (1981) and Bénabou and Pycia (2002), *cognitive system* versus *affective system* in Brocas and Carillo (2008a). The hierarchical and modular structure described in Minsky (1986), Oatley (1986), and Blakeslee (2005) are also in line with our model. More recently, Alonso, Brocas, and Carillo (2011) propose a *principal–agent* model of brain where a coordinator (i.e., a principal) allocates limited resources (e.g., oxygen, glucose) to brain systems responsible for different tasks. The *principal* in our model has a higher position in the hierarchy compared to the *agents* and deals with a more complex problem of optimal task assignment. We denote the risk-neutral principal as P and agents as A_i, $i = 1, 2$.

States of the World and Specialization

In our model there are signals that partition the information space, that is, describe the states of the world. States of the world are characteristics of the world.[5] For instance, they can be interpreted as different aspects or characteristics of a unique physical state of the world. For simplicity, we assume that there are two pure characteristics of the world, denoted by $S_j \in \{S_1, S_2\}$ that can be described by signals. Signals are defined on the power set of these characteristics.

We denote signals by s_j, where originally $s_j \in \{s_1, s_2, s_3, s_4\}$. The possible signals are as follows:

$$\begin{aligned}
&s = s_1 \rightarrow \{s_1\}\, x S_2, \\
&s = s_2 \rightarrow S_1 x \{s_2\}, \\
&s = s_3 = \{s_1, s_2\} \rightarrow \{s_1\}\, x \{s_2\}, \text{ and} \\
&s = s_4 = \emptyset \rightarrow S_1 x S_2
\end{aligned}$$

where the first case corresponds to receiving a signal that gives information only about S_1, the second one corresponds to receiving a signal that gives information only about S_2, the third one implies that the signal received gives information about both S_1 and S_2, and the last one implies that the signal received gives information about neither S_1 nor S_2.

[4] An alternative theory of mind (the so called *domain-general processing theory of mind*) stipulates that mind is not modular, the mental activity is distributed across mind in a complex fashion and cannot be decomposed into separate units. William Uttal is the major proponent of this approach (see Uttal, 2002; Uttal, 2003). In this article, we adopt the modular mind approach proposed by Fodor (1983) and Minsky (1986).

[5] Therefore, we do not use the term *state of the world* to refer to different virtual states of the world.

Each agent A_i ($i = 1, 2$) has a comparative advantage in perceiving one pure signal s_j, where $j = 1, 2$. The specialization (or the comparative advantage) is associated with agents' perception quality. Agent A_i perceives the signal that he is specialized in with probability $\bar{p}$ and the signal that he is not specialized in with probability $\underline{P}$, where $\bar{P} + \underline{P} = 1$ and $\bar{P} > \underline{P}$. These modeling assumptions are also compatible with some findings in neuroscience, which stipulate that different brain systems perform different and sometimes incompatible tasks and a selection should be made among competing systems (see Zeki et al., 1991; Watson et al., 1993; Kanwisher et al., 1997; Berridge and Robinson, 2003; Dityatev et al., 2010).

In each period t, a signal arrives P, which is then assigned by P to A_1, A_2 or both. By assigning the signal, the principal uses one (or both) of the agents in the decision-making. Here, we do not model the task explicitly: it can be interpreted as making a simple decision given the signal in the current period. If the signal is successfully perceived, the task will be performed successfully, which brings a finite, real-valued return of π (or $\Pi > \pi$, in the case of a composite signal) to the individual; if the agent who is delegated cannot perceive the signal successfully, then he cannot perform the task successfully, which leads to 0 return for the individual. Returns can be interpreted as any kind of benefits from the successful completion of the task. They can be material returns or utilities in general.

Cognitive Dissonance

The cognitive dissonance comes into the picture in the case of a composite signal, s_3. The composite signal describes two different aspects of the world, which in turn requires both agents to act together. Without both of them performing, the task in s_3 cannot be successfully completed. However, because these two agents are in conflict with each other, their simultaneous involvement creates a cognitive dissonance, which influences individual returns negatively (see Baumeister, 2003; Elster, 2004 for a similar phenomena between cognitive system and affective system).[6] In our model, cognitive dissonance that is due to agent A_i is denoted as θ_i. θ_i ($i = 1, 2$) takes values in $[0, \Theta]$.

While we are describing this modular structure found in scientific research, we would also like to mention the possibility of conflict among different parts of the brain. Conflicting interests of different parts of the mind/brain is also a well-documented phenomenon in neuroscience and psychology. Among all, Mischel et al. (1989), Tversky and Shafir (1992), McClure et al. (2004), Bechara (2005), and Camerer et al. (2005) are some studies that report on the existence of a conflict between different mental and neural systems. In a theoretical study, Livnat and Pippenger (2006) argue that optimal brain system, which is designed for a sole purpose, can have agents that are in conflict with each other. The conflict reported in these studies is one of the cornerstones of our model. It is the existence of this conflict that leads to cognitive dissonance.

Costs of Intrapersonal Conflict

A phenomenon, which is directly related to conflict and cognitive dissonance, is the tension and the corresponding fall in performance. Conflicting cognitions have the potential to prevent individual from behaving effectively (see Jones and Gerard, 1967; Spadafore, 1976; Miller and Cohen, 2001; Baumeister, 2003; Harmon-Jones, Amodio and Harmon-Jones, 2009; Gawronski, 2012). This modeling assumption is also in conformity with the *action-based model* of cognitive dissonance, which suggests that the presence of cognitive dissonance may lead to ineffective and/or conflicted actions (see Harmon-Jones, Amodio and Harmon-Jones, 2009; Gawronski, 2012; Harmon-Jones, 2012, and references therein) and hence a fall in performance.

A simple and useful example comes from a well-known psychological test, the *Stroop test*

[6] The cognitive dissonance that can be experienced as a result of the conflict of decisions made in different periods is not the focus of this article. Gur and Sackheim's (1979) definition of *self-deception* as a motivated act is an appropriate explanation for this phenomenon. The fact that some brain areas (including consciousness) may be unaware of information to which other brain areas respond is a well-established phenomenon in neuroscience (see Berns et al., 1997 and Whalen et al., 1998). This kind of cognitive dissonance may also be reduced by what is called in the literature as *motivated forgetting* or *selective memory.*

named after John Ridley Stroop: Blakeslee (2005) analyzes the nature of hypnosis and suggestion on subjects' decision-making performance in a psychological experiment. Subjects are shown cards on which there are colored writings and asked to press a button to tell the color of the writing. The trick in the experiment is that writings conflict with the colors; for example, the word "blue" is written in red or "yellow" is written in green. In this experiment, it has been observed that subjects sometimes spend some time to tell the color and sometimes even tell the wrong color. Apparently, the color signal and the word signal create a conflict in the brain. This is called the *Stroop effect.* She reports that hypnosis and suggestion stop this conflict by suppressing the parts of the brain that is responsible for reading and detecting conflict. Aside from the hypnosis part, the conflict and the *Stroop effect* in her study are very much related to the structure of our model. The very existence of conflict among different cognitions has a potential to reduce the task performance or decision quality.

In our model, the corresponding fall in returns attributable to cognitive dissonance of θ_i is denoted as $\phi_i(\theta_i)$. We assume that $\phi_i(.)$ is a continuous, increasing, and concave function. We normalize $\phi_i(0)$ to be 0. Moreover, $\phi_i(\Theta) = \overline{\Theta}$.

Before we describe the cross-training phase and the value function, we present the timeline of events below (see Figure 1).

Habituation and Cross-Training Attempts

The motivation for incorporating the *cross-training phase* into the model comes from a psychological phenomenon called *habituation.* The term *habituation* in the psychology literature refers to the decline of a conditioned response following repeated exposure to the conditioned stimulus. Some examples would be as follows:

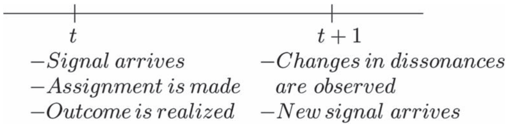

Figure 1. Timeline of events.

- When a subject in an experiment is given an electric shock repetitively with the same magnitude, the experimenter needs to increase the magnitude of the shock to get a response that is equivalent to the one at the beginning.
- Old people who react to social changes, fashion, and the habits of a new generation start to get used to them if they are frequently exposed to these phenomena.
- An individual who thinks that lying is not a good behavioral trait may feel less dissonance if he occasionally has to lie in his new job.

Habituation is a type of defense mechanism for the individual to avoid the unpleasant feeling he or she experiences. The reason is the tension that occurs in the case of dissonance. If a particular action/belief/cognition is in contradiction with some other and is comparatively more crucial for the individual, he or she needs to adjust the latter one accordingly so that he or she feels less dissonance. Habituation may take place consciously or unconsciously.

Habituation is an important part of our model because it creates a trade-off for the individual between current and future returns, making principal's decision problem a dynamic one. The effect of habituation is observed on dissonance levels, that is, θ_i, $i = 1, 2$. θ_is are *discrete, one-to-one,* and *bounded* functions of successful crosstraining attempts in the model. If an agent is loaded with the signal that he is not specialized in (i.e., cross-trained) and if he perceives the signal successfully, it decreases his cognitive dissonance that he experiences in the composite state. Hence, for all i, $\frac{\Delta\theta_i}{\Delta n_i} < 0$ at each point (where n_i is the number of successful cross-training attempts) except when $\theta_i = 0$. Moreover, we assume that the zero lower bound can be reached with finitely many successful cross-training attempts.

In the composite state, the principal needs both agents to be able to perform the task successfully and when both agents act together in the same time period on the same task, they experience cognitive dissonance attributable to their awareness of the "other," which otherwise does not happen because of the information asymmetries prevalent (see Berns et al., 1997;

Table 1
Immediate Returns

	A_i specialized in s_1	A_i specialized in s_2	Both Agents
s_1	$\underline{P}.\pi + (1 - \underline{P}).0$	$\bar{P}.\pi + (1 - \bar{P}).0$	*NA*
s_2	$\underline{P}.\pi + (1 - \underline{P}).0$	$\bar{P}.\pi + (1 - \bar{P}).0$	*NA*
s_3	0	0	$((\bar{P}\,\bar{P}) + (\underline{P}\underline{P})).\Pi - (\bar{P}\bar{P})[\phi_1(\theta_1) + \phi_2(\theta_2)]$

Rauch et al., 1997; Knowlton et al., 1996 for neuroscientific evidence for asymmetric information in the brain and Fodor, 1983 for *information encapsulation,* resembling asymmetric information, as a necessity for a modular mind). We assume that cognitive dissonance is experienced only in the case of successful completion of the task, which requires perception of both parts of the composite signal. In a sense, successful perception of the signal by one agent guarantees the other agent's awareness of the agent's active involvement.

The possibility of reducing cognitive dissonance in composite states raises a question for the principal: cross-training or acting in accordance with agents' specializations? The probability of receiving signal s_j, where $j = 1, 2, 3$ is independent and identical across time and equal to $F_j = 1/3$, for all $j = 1, 2, 3$.[7] Therefore, the principal should decide whether to act in accordance with specializations or attempt cross-training, when he receives the signal s_j, $j = 1, 2$.

It is necessary—for clarification—to make a distinction between the cognitive dissonance experienced when the agent attempts cross-training and the cognitive dissonance experienced in the presence of a composite signal. The former one is endogenous, that is, cognitive dissonance is intentionally chosen to be experienced, whereas the latter is exogenous, that is, not chosen by the agent. In the former, the negative effect of dissonance takes the form of a fall in perception quality, whereas in the latter it takes the form of an explicit fall in the return.[8]

Motivated habituation in our model has an investment character, and examples for such behavior can be observed in real life instances in which an individual–anticipating future dissonance-inducing states of the world and a corresponding fall in returns he is likely to experience–tries to reduce the dissonance he will have by exposing his split selves to dissonance-inducing signals.

Having described the important notions in the model, below we define the individual's value function.

The Value Function

The expected current returns levels under different signals and assignments are presented below in Table 1 for the reader to be able to follow easily. For instance, the first cell can be read as the expected return if s_1 is assigned to A_i, who is specialized in s_1.[9]

Now, we can write the principal's value function. The principal's problem at each period is *to determine which agent(s) he should assign the signal/task in the given state:*

$$V(j, \theta_1, \theta_2) = \max \left\{ \begin{array}{l} \bar{P}\pi + \beta EV_{\max}(j', \theta_1', \theta_2'), \\ \underline{P}\pi + \beta EV_{\min}(j', \theta_1', \theta_2') \end{array} \right\}, \quad \text{if } j = 1, 2 \quad (1)$$

[7] Because the signal that does not describe any characteristic of the world (i.e., $s_4 = \emptyset$) is not operational in any sense, we assumed $F_4 = 0$.

[8] I thank an anonymous reviewer for pointing out this important distinction.

[9] We assume that for a singleton signal such as s_j, $j = 1, 2$, the decision is between *deciding in accordance with specializations* and *attempting cross-training.* This is more of a simplifying assumption to keep the model tractable. Neither will it bring additional insights nor would our main result qualitatively change (only the cross-training will take shorter time) if we relax this assumption.

$$V(j,\theta_1,\theta_2) = \max\left\{\begin{array}{c}(\overline{P}\overline{P})(\Pi - [\phi_1(\theta_1)+\phi_2(\theta_2)]) + \beta EV_{NN}(j',\theta_1',\theta_2'),\\ 0+\beta EV_{YN}(j',\theta_1',\theta_2'),\\ 0+\beta EV_{YN}(j',\theta_1',\theta_2'),\\ (\underline{P}\underline{P})(\Pi - [\phi_1(\theta_1)+\phi_2(\theta_2)]) + \beta EV_{YY}(j',\theta_1',\theta_2')\end{array}\right\},\ \text{if } j=3 \quad (2)$$
$$\text{given } \theta_i(n_i) \text{ for } i = 1,\ 2,$$

where $V_{\max}$ refers to next period's value function when the task in the current state j is assigned to the agent specialized in state j and $V_{\min}$ refers to next period's value function when the task in the current state j is assigned to the agent not specialized in state j. When $j = 3$, there are four different options for the principal:

- assigning in accordance with specializations for both signals,
- cross-training A_1 and not cross-training A_2,
- not cross-training A_1 and cross-training A_2, and
- cross-training both A_1 and A_2.

Subscripts *NN, YN, NY,* and *YY* correspond to these options, respectively. In *YN* and *NY*, because one of the signals is not assigned, it will not be perceived for sure. Therefore, the current return is zero. Expectations in the next period value functions are taken on j and θ_is. The reason for taking expectations on θ_is is the stochastic perceptions of agents. Other than that, θ_is follow a known deterministic process (i.e., function of number of successful cross-training attempts) and they take values in $[0, \Theta]$ interval. Also, note that n_is need not be state variables because, given $\theta_i(n_i)$ functions, once θ_i is observed the principal knows the value of $\theta_i(n_i + 1)$. The fact that $\theta_i(n_i)$ are one-to-one functions is useful here.

The Optimal Solution

The two lemmas below provide some preliminary results to be used in the main theorem. The first lemma suggests that there exists a certain time period, after which there is no expected future benefits from cross-training, which implies that the principal decides solely based on agents' specializations.

Lemma 1

The probability that there exists a finite period $T \geq 1$ after which "acting in accordance with specializations" is the unique optimal action for all $t \geq T$ is 1.

Proof

See Appendix.

The lemma below suggests that there is no reason to wait for cross-training if there are expected future benefits from doing so.

Lemma 2

It is not optimal to attempt to cross-train A_i in the future, if it is not optimal now.

Proof

See Appendix.

The following theorem presents our main result. It encapsulates that the optimal decision rule for the principal has a threshold structure.

Theorem 1

Given the stochastic dynamic programming problem in (1) and (2), the optimal decision rule has the following threshold structure with crucial values of θ_i, that is, $\bar{\theta}_i$, for $i = 1, 2$:

- Attempt cross-training in state $j = 1, 2$ if $\theta_i \geq \bar{\theta}_i$ where agent i is the one, not specialized in state j and
- Act in accordance with specializations if $\theta_i < \bar{\theta}_i$.

The difficulty of solving the stochastic dynamic programming problem given above stems from the fact that the evolution of the stochastic processes in the future depends on the current decision. Therefore, it is not a stationary stochastic dynamic programming problem that can be solved easily by equating the expected returns that are due to different decisions to each other. However, Lemma 1 and Lemma 2 stated above are useful in proving this result. The facts we exploit are (i) after finitely many periods, the problem becomes a stationary one with probability 1 (see Lemma 1) and (ii) postponing cross-training when it is possible in the current period is not an optimal action (Lemma 2).

Proof

See Appendix.

The intuition for this result comes from a simple cost–benefit analysis. Basically, what the principal does is compare the expected costs and benefits of each action at each point in time. We show that there will be a finite time period after which incurring the (expected) current costs of cross-training is not expected to pay sufficiently in the future. This is determined by the level of cognitive dissonances and the process governing them along with other parameters of the model. When crosstraining becomes an unattractive option, the principal starts to make decisions solely based on specializations of the agents.

The following corollary, which presents the main message of the article, follows from Theorem 1.

Corollary 1

Cognitive dissonance (intrapersonal conflict) can be a long-run phenomenon in an optimizing mind.

Proof

Directly follows from Theorem 1.

If it is not worth reducing the intrapersonal conflict to zero, the individual continues to feel the tension in every instance where he needs to use both selves in a decision-making problem. Our result is consistent with the findings of Williams and Aaker (2002), who experimentally show that mixed/conflicting emotions may peacefully coexist (especially in individuals with higher propensity to accept duality) and Livnat and Pippenger (2006), who theoretically show that an optimal brain can be composed of conflicting parts. Our analytical result is also consistent with Gawronski and Kulakowski (2007), who computationally show that the complete removal of cognitive dissonance can take arbitrarily long time.

The following corollaries present results on some *special cases* in which either one or both of the agents are completely resistant to habituation.

Corollary 2

(Fixed θ_1 and θ_2) If both agents are totally resistant to habituation, that is, their dissonance terms are constant, the principal does not attempt cross-training and decides according to the specialization of agents.

Proof

See Appendix.

This kind of situation may be observed if an individual has very strong beliefs (or in general cognitions), which are in conflict with each other.

Corollary 3

(Fixed θ_1 (θ_2) and variable θ_2 (θ_1)) If agent A_i is totally resistant to habituation, that is, his dissonance term, $\theta_i(n_i) = \theta_i(0)\ \forall n_i \in \mathbb{N}$, where n_i is the number of successful cross-training attempts, then the principal does not attempt to crosstrain A_i. This corresponds to assigning the task in the state A_i is not specialized in to agent $A_{k\neq i}$. On the other hand, if agent $k \neq i$ is flexible (i.e., not totally resistant to habituation) the optimal decision rule in Proposition 1 is valid for A_k, that is, the principal assigns the task in state $A_{i\neq k}$ is specialized in to A_k until θ_k reaches $\bar{\theta}_k$ and then starts to assign the task to A_i.

Proof

See Appendix.

This kind of situation may be observed if an individual has conflicting cognitions, one of which is very strong and resistant whereas the other is not.

Comparative Static Analyses

In this section, we conduct some comparative static analyses and look at the impact of a change in some important parameters of the model on the optimal (long run) values of dissonances. These parameters are the probability of the composite state, the pace of habituation, and the discount factor. We derive intuitive implications from these analyses and give examples from real life.

Proposition 1

An increase in the probability of the composite signal, F_3, decreases the long run levels of cognitive dissonances, that is, $\tilde{F}_3 > F_3 \Rightarrow \tilde{\bar{\theta}}_i < \bar{\theta}_i \; i = 1, 2$.

Proof

See Appendix.

This is a very intuitive result that shares a flavor of what we also observe in our daily lives. All it says is that people are less willing to forego current returns by investing in something that is not expected to pay back very frequently in the future; or equivalently, an increase in the frequency of future returns (holding returns per-period fixed) makes the investment more attractive. The frequency of expected future returns has a major impact on the decision to bear the current costs of an investment project. This comparative static result is also parallel to the effect of interaction frequency on balance in Heiderian theory of balance.

Proposition 2

An increase in the rate of habituation (increase in $\left|\frac{\Delta\theta_i}{\Delta n_i}\right|$ at every point) decreases the long run levels of cognitive dissonances, that is, $\tilde{\theta}_i(n_i) \le \theta_i(n_i)$ and $\frac{\Delta\tilde{\theta}_i}{\Delta n_i} < \frac{\Delta\theta_i}{\Delta n_i} \Rightarrow \tilde{\bar{\theta}}_i < \bar{\theta}_i, \; i = 1, 2$.

Proof

See Appendix.

The pace of habituation directly influences the total (opportunity) cost of the investment to be made to achieve the minimal dissonance level. If the pace of habituation is very rapid, the principal can achieve the minimum level of dissonance in a few cross-training attempts, which makes the opportunity cost very small and hence the investment very attractive. On the other hand, if the pace of habituation is very slow, it will take so many attempts to achieve the minimum level of dissonance, which leads to a very high opportunity cost and makes the investment unattractive.

Note that the increase in habituation rate is the same for all A_i. If the magnitude of change is assumed to be different among i's, qualitatively the same result would hold. However, the long run level of conflict would be different depending on the relative flexibility of A_i's. The flexibility/rigidity, in a sense, is related to the easiness (or difficulty) of changing/relaxing/giving up beliefs, traits, habits and so forth As the model implies, the more flexible the split-self, easier to cross-train him, which leads to a lower dissonance on his side in the long run.

Proposition 3

An increase in the discount factor β decreases the long run levels of cognitive dissonances, that is, $\tilde{\beta} > \beta \Rightarrow \tilde{\bar{\theta}}_i < \bar{\theta}_i \; i = 1, 2$.

Proof

See Appendix.

The logic behind this result is very intuitive. An increase in patience, implied by an increase in β, means a higher valuation for future benefits, which can be obtained by cross-training. To grasp this result clearly, one may think of two extreme type of individuals (i.e., principals in our context) one of which is completely myopic (i.e., with $\beta = 0$) and the other is perfectly forward looking (i.e., with $\beta = 1$). The first one cares only about *today*, and the latter value returns in every period equally. The first one, when faced with a decision after receiving a signal s_j where $j = 1, 2$, assigns the task to the agent specialized in that signal, and the latter may attempt cross-training depending on the value of θ_i and other parameters.

Conclusion

We present a dynamic model of individual decision-making problem under cognitive dissonance. Our model also incorporates well-documented cognitive and psychological phenomena such as habituation, modularity of and hierarchy in the mind. It proposes *motivated habituation* as an alternative method for avoiding cognitive dissonance and identifies conditions under which an optimizing mind seeks to engage in costly cognitive dissonance reduction. We conclude that under certain circumstances, an optimizing mind does not seek dissonance reduction. Specifically, an individual may not find it worthwhile to reduce the intrapersonal conflict he experiences in some states of the world. This may be attributable to factors such as the specialization levels of split-selves in different signals, resistance/openness to habituation, frequency of the conflict inducing state, and time preferences (i.e., forward-lookingness). Our main result is in line with some earlier findings in the psychology literature:

- An optimizing brain may consist of conflicting agents (see Livnat and Pippenger, 2006),
- The presence of cognitive dissonance does not necessarily imply irrational behavior (see Lévy-Garboua and Blondel, 2002),
- It may take arbitrarily long to remove all cognitive dissonance (see Gawronski and Kulakowski, 2007),
- Living with dissonance is a resolution strategy people sometimes use (see Mahaffy, 1996)
- A person may continue to maintain conflicting attitudes and incur the costs of this tension (see Harmon-Jones et al., 2009), and
- Conflicting emotions may coexist, especially more in individuals who are more open to duality (see Williams and Aaker, 2002).

We also conduct some comparative static analyses on model parameters, which give useful insights. In particular, an increase in the pace of habituation, the discount parameter, and the frequency of conflict have the same kind of impact on long run levels of dissonances. An increase in any of these parameters reduces the optimal level of dissonance, increasing the expected benefit from attempting cross-training.

Our model can also be interpreted in a group/team context with interpersonal conflict (see Heider, 1958; Spadafore, 1976; Cooper and Mackie, 1983; Matz and Wood, 2005; Glasford et al., 2008). In that context, the principal in our model can be thought of as a manager who is responsible for maximizing the discounted sum of returns by allocating randomly arriving tasks to individuals in his team or the team as a whole. The individuals have dissonance-inducing characteristics (e.g., different working methods, specializations, etc.), which lead to a fall in team performance under some circumstances. We theorize that our main results would carry on to this setup, as well.

Before concluding, we want to emphasize that our model does not aim to explain the emergence of cognitive dissonance, but rather takes cognitive dissonance as given and focuses on optimal dissonance reduction behavior. Obviously, explaining the emergence of cognitive dissonance endogenously in a mathematical framework is a promising venue for future research. It is also worth mentioning that our model framework applies more to the long-term, persistent cognitive dissonance, whereas short-term dissonance could be removed in reality. Future research may use optimization approach along with other mathematical techniques (e.g., graph theory) to investigate dissonance reduction. In the context of our model, the study of the implications of a time-inconsistent principal and the need for cognitive consistency in a temptation versus self-control problems would also be fruitful. Finally, we believe that an experimental test of the motivated habituation argument is possible and would be of interest.

References

Akerlof, G. A., & Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *American Economic Review, 72,* 307–319.

Alonso, R., Brocas, I., & Carillo, J. D. (2011). Resource allocation in the brain. *CEPR Discussion Paper #* 8408.

Aronson, E. (1998). *The social animal*. New York, NY: Freeman.

Badre, D., & D'Esposito, M. (2007). FMRI evidence for a hierarchical organization of the prefrontal cortex. *Journal of Cognitive Neuroscience, 19,* 1–18.

Balestrino, A., & Ciardi, C. (2008). Social norms, cognitive dissonance and timing of marriage. *Journal of Socio-Economics, 37,* 2399–2410.

Baumeister, R. (2003). The psychology of irrationality: Why people make foolish, self-defeating choices. In I. Brocas & J. D. Carrillo (Eds.), *The psychology of economic decisions. Vol. 1: Rationality and well-being* (pp. 3–16). Oxford, UK: Oxford University Press.

Beauvois, J.-L., & Joule, R.-V. (1996). *A radical theory of dissonance*. New York, NY: Taylor and Francis.

Bechara, A. (2005). Decision making, impulse control and loss of willpower to resist drugs: A neurocognitive perspective. *Nature Neuroscience, 8,* 1458–1463.

Bénabou, R., & Pycia, M. (2002). Dynamic inconsistency and self-control: A planner-doer interpretation. *Economics Letters, 77,* 419–424.

Bendersky, C., & Curhan, J. R. (2009). Cognitive dissonance in negotiation: Free choice or justification. *Social Cognition, 27,* 455–474.

Berns, G., Cohen, D., & Mintun, M. (1997). Brain regions responsive to novelty in the absence of awareness. *Science, 276,* 1272–1275.

Berridge, K. C., & Robinson, T. E. (2003). Parsing reward. *Trends in Neurosciences, 26,* 507–513.

Bisin, A., & Iantchev, E. (2010). *Evolutionary selection of modular decision architectures*. Unpublished manuscript.

Blakeslee, S. (2005). This is your brain under hypnosis. *The New York Times - Science,* November 22, 2005.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108,* 624–652.

Brady, G. L., Clark, J. R., & Davis, W. L. (1995). The political economy of dissonance. *Public Choice, 82,* 37–51.

Brocas, I., & Carillo, J. D. (2008a). The brain as a hierarchical organization. *American Economic Review, 98,* 1312–1346.

Brocas, I., & Carillo, J. D. (2008b). Theories of the mind. *American Economic Review Papers and Proceedings, 98,* 175–180.

Camerer, C., Loewenstein, G., & Prelec, D. (2005). Neuroeconomics: How neuroscience can inform economics. *Journal of Economic Literature, 43,* 9–65.

Carruthers, P. (2006). *The architecture of the mind*. Oxford, UK: Oxford University Press.

Charman, T., & Baron-Cohen, S. (1995). Understanding photos, models, and beliefs: A test of the modularity thesis of theory of mind. *Cognitive Development, 10,* 287–298.

Cooper, J., & Mackie, D. (1983). Cognitive dissonance in an intergroup context. *Journal of Personality and Social Psychology, 44,* 536–544.

Cosmides, L., & Tooby, J. (1987). From evolution to behavior: Evolutionary psychology as the missing link. In J. Dupre (Ed.), *The latest on the best: Essays on evolution and optimality*. Cambridge, MA: The MIT Press.

Cosmides, L., & Tooby, J. (1992). The psychological foundations of culture. In J. Barkow, L. Cosmides, & J. Tooby (Eds.), *The adapted mind: Evolutionary psychology and the generation of culture*. Oxford, UK: Oxford University Press.

Dickens, W. T. (1986). Crime and punishment again: The economic approach with a psychological twist. *Journal of Public Economics, 30,* 97–107.

Dickinson, D., & Oxoby, R. J. (2011). Cognitive dissonance, pessimism, and behavioral spillover effects. *Journal of Economic Psychology, 32,* 295–306.

Dityatev, A., Schachner, M., & Sonderegger, P. (2010). The dual role of the extracellular matrix in synaptic plasticity and homeostasis. *Nature Reviews Neuroscience, 11,* 735–746.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in Neurosciences, 23,* 475–483.

Elster, J. (2004). Costs and constraints in the economy of the mind. In I. Brocas & J. D. Carillo (Eds.), *The psychology of economic decisions, Vol. 2: Reasons and choices* (pp. 3–14). Oxford, UK: Oxford University Press.

Epstein, L. G., & Kopylov, I. (2006). Cognitive dissonance and choice. Rochester Center for Economic Research Working Paper #525.

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Fischbacher, U., & Heusi, F. (2008). Lies in disguise. An experimental study on cheating. *Thurgau Institute of Economics Research Paper #* 40.

Fodor, J. (1983). *The modularity of the mind*. Cambridge, MA: Bradford Books, MIT Press.

Garvey, B. (2012). Free will, compatibilism, and the human nature wars: Should we be worried? *Philosophy of Behavioral Biology, Boston Studies in the Philosophy of Science, 282,* 255–274.

Gawronski, B. (2012). Back to the future of dissonance theory: Cognitive consistency as a core motive. *Social Cognition, 30,* 652–668.

Gawronski, P., & Kulakowski, K. (2007). A numerical trip to social psychology: Long-living states of cognitive dissonance. *Physica-ph,* 2550–2558.

Geary, D. C., & Huffman, K. J. (2002). Brain and cognitive evolution: Forms of modularity and functions of mind. *Psychological Bulletin, 128,* 667–698.

Gilad, B., Kaish, S., & Loeb, P. D. (1987). Cognitive dissonance and utility maximization. *Journal of Economic Behavior and Organization, 8,* 61–73.

Gino, F., Ayal, S., & Ariely, D. (2009). Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. *Psychological Science, 20,* 393–398.

Gino, F., Norton, M., & Ariely, D. (2010). The counterfeit self: The deceptive costs of faking it. *Psychological Science, 21,* 712–720.

Glasford, D. E., Pratto, F., & Dovidio, J. F. (2008). Intragroup dissonance: Responses to ingroup violation of personal values. *Journal of Experimental Social Psychology, 44,* 1057–1064.

Goldsmith, A. H., Sedo, S., Darity, W., & Hamilton, D. (2004). The labor supply consequences of perceptions of employer discrimination during search and on the-job: Integrating neoclassical theory and cognitive dissonance. *Journal of Economic Psychology, 25,* 15–39.

Gur, R. C., & Sackheim, H. A. (1979). Self-deception: A concept in search of a phenomenon. *Journal of Personality and Social Psychology, 37,* 147–169.

Harmon-Jones, E. (1999). Toward an understanding of the motivation underlying dissonance effects: Is the production of aversive consequences necessary to cause dissonance? In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology* (p. 71–99). Washington D. C.: American Psychological Association.

Harmon-Jones, E. (2012). Cognitive dissonance theory. In V. S. Ramachandran (Ed.), *The encyclopedia of human behavior* (Vol. 1, pp. 543–549). Burlington, VT: Academic Press.

Harmon-Jones, E., Amodio, A. M., & Harmon-Jones, C. (2009). Action-based model of dissonance: A review, integration, and expansion of conceptions of cognitive conflict. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 119–166). Burlington, VT: Academic Press.

Heider, F. (1946). Attitudes and cognitive organization. *Journal of Psychology, 21,* 107–112.

Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.

Hughlings-Jackson, J. (1959). *Selected writings of John Hughlings-Jackson* (J. Taylor, Ed.). New York, NY: Basic Books.

Ishida, J. (2010). Vision and flexibility in a model of cognitive dissonance. The Institute of Social and Economic Research Discussion Paper # 771.

Johansson-Stenman, O., & Svedsäter, H. (2008). Measuring hypothetical bias in choice experiments: The importance of cognitive consistency. *The B. E. Journal of Economic Analysis and Policy, 8,* Article 41.

Jones, E. E. (1985). Major developments in social psychology during the past five decades. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (3rd ed., pp. 47–108). New York, NY: Random House.

Jones, E. E., & Gerard, H. B. (1967). *Fundamentals of social psychology*. New York, NY: Wiley.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience, 17,* 4302–4311.

Knowlton, B. J., Mangels, J. A., & Squire, L. R. (1996). A neostriatal habit learning system in humans. *Science, 273,* 1399–1402.

Konow, J. (2000). Fair shares: Accountability and cognitive dissonance in allocation decisions. *American Economic Review, 90,* 1072–1091.

Lester, D., & Yang, B. (2009). Two sources of human irrationality: Cognitive dissonance and brain dysfunction. *Journal of Socio-Economics, 38,* 658–662.

Lévy-Garboua, L., & Blondel, S. (2002). On the rationality of cognitive dissonance. In S. Grossbard-Schechtman & C. Clague (Eds.), *The expansion of economics: Towards a more inclusive social science*. New York, NY: M E Sharpe Inc.

Livnat, A., & Pippenger, N. (2006). An optimal brain can be composed of conflicting agents. *Proceedings of the National Academy of Sciences of the United States of America, 103,* 3198–3202.

Mahaffy, K. M. (1996). Cognitive dissonance and its resolution: A study of lesbian Christians. *Journal for the Scientific Study of Religion, 35,* 392–402.

Margolin, L. (1997). *Under the cover of kindness: The invention of social work*. Charlottesville, VA: The University Press of Virginia.

Matthey, A., & Regner, T. (2011). Do I really want to know? A cognitive dissonance-based explanation of other-regarding behavior. *Games, 2,* 114–135.

Matz, D. C., & Wood, W. (2005). Cognitive dissonance in groups: The consequences of disagreement. *Journal of Personality and Social Psychology, 88,* 22–37.

McClure, S., Laibson, D., Loewenstein, G., & Cohen, J. (2004). Separate neural systems value immediate and delayed monetary rewards. *Science, 306,* 503–507.

Miller, B. L., & Cummings, J. L. (2007). *The human frontal lobes, 2nd ed.: Functions and disorders (The science and practice of neuropsychology)*. New York, NY: The Guilford Press.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202.

Minsky, M. (1986). *The society of mind*. New York, NY: Simon & Schuster.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science, 244,* 933–938.

Oatley, K. (1986). Conflict and control among mental agents. *Journal of Semantics, 5,* 165–168.

Oxoby, R. J. (2003). Attitudes and allocations: Status, cognitive dissonance, and the manipulation of attitudes. *Journal of Economic Behavior and Organization, 52,* 365–385.

Pinker, S. (1997). *How the mind works*. New York, NY: Norton and Company.

Poldrack, R., & Rodriguez, P. (2004). How do memory systems interact? Evidence from human classification learning. *Neurobiology of Learning and Memory, 82,* 324–332.

Rabin, M. (1994). Cognitive dissonance and social change. *Journal of Economic Behavior and Organization, 23,* 177–194.

Rauch, S. L., Whalen, P. J., Savage, C. R., Curran, T., Kendrick, A., Brown, H. D., . . . Rosen, B. R. (1997). Striatal recruitment during an implicit sequence learning task as measured by fMRI. *Human Brain Mapping, 5,* 124–132.

Sakai, H. (1999). A multiplicative power-function model of cognitive dissonance: Toward an integrated theory of cognition, emotion, and behavior after Leon Festinger. In E. Harmon-Jones & J. Mills (Eds.), *Cognitive dissonance: Progress on a pivotal theory in social psychology, science conference series* (pp. 267–294). Washington, DC: American Psychological Association.

Sakai, H. (2013). A recurrent network model of cognitive consistency: How balance and dissonance theories could be unified? *The Sapporo University Journal, 35,* 43–55.

Samuels, R. (2006). Is the human mind massively modular? In R. Stainton (Ed.), *Contemporary debates in cognitive science* (pp. 37–56). Oxford, UK: Blackwell.

Sergent, J., & Signoret, J. L. (1992). Varieties of functional deficits in prosopagnosia. *Cerebral Cortex, 2,* 375–388.

Shettleworth, S. J. (1998). *Cognition, evolution and behavior*. New York, NY: Oxford University Press.

Shultz, T. R., & Lepper, M. R. (1996). Cognitive dissonance reduction as constraint satisfaction. *Pscyhological Review, 103,* 219–240.

Smith, J. (2009). Cognitive dissonance and the overtaking anomaly: Psychology in the principal-agent relationship. *Journal of Socio-Economics, 38,* 684–690.

Spadafore, G. J. (1976). Cognitive dissonance paradigm: A technique to improve the effectiveness of the team approach. *Psychological Reports, 38,* 1295–1299.

Sperber, D. (1994). The modularity of thought and the epidemiology of representations. In L. A. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 39–67). New York, NY: Cambridge University Press.

Thaler, R., & Shefrin, H. M. (1981). An economic theory of self-control. *Journal of Political Economy, 89,* 392–406.

Tversky, A., & Shafir, E. (1992). The disjunction effect in choice under uncertainty. *Psychological Science, 3,* 358–361.

Uttal, W. (2002). Précis of the new phrenology: The limits of localizing cognitive processes in the brain. *Brain and Mind, 3,* 221–228.

Uttal, W. (2003). *The new phrenology: The limits of localizing cognitive processes in the brain*. Cambridge, MA: MIT Press.

Watson, J. D. G., Myers, R., Frackowiak, R. S. J., Hajnal, J. V., Woods, R. P., Mazziotta, J. C., . . . Zeki, S. (1993). Area V5 of the human brain: Evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex, 3,* 79–94.

Whalen, P., Rauch, S., Etcoff, N., McInerney, S., Lee, M., & Jenike, M. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *The Journal of Neuroscience, 18,* 411–418.

Williams, P., & Aaker, J. L. (2002). Can mixed emotions peacefully coexist? *Journal of Consumer Research, 28,* 636–649.

Zeki, S., Watson, J. D., Lueck, C. J., Friston, K. J., Kennard, C., & Frackowiak, R. S. (1991). A direct demonstration of functional specialization in human visual cortex. *The Journal of Neuroscience, 11,* 641–649.

Zhou, H. (2000). Working with a dying ideology: Dissonance and its reduction in Chinese journalism. *Journalism Studies, 1,* 599–616.

(*Appendix follows*)

Appendix

Proofs

Proof of Lemma 1

The proof directly follows from the following facts: (i) θ_i for $i = 1, 2$ is bounded from below by 0, (ii) Probability of receiving s_j is independent across time and equals to $\frac{1}{3}$ for all $j = 1, 2, 3$, (iii) $\underline{P} > 0$, and (iv) if θ_i reaches 0, then it reaches in finite number of periods with probability 1, that is, n_i that satisfies $\theta_i(n_i) = 0$ is finite with probability 1. After finitely many periods with successful cross-training, cross-training will not increase expected future returns, after which the principal will decide in accordance with specializations of agents.

Proof of Lemma 2

The proof follows mainly from the fact that θ_is are decreasing functions of successful cross-training attempts (n_i), there is no reversal in θ_is, and ϕ_is are increasing functions of dissonances (θ_is). Suppose that at time t, a signal s_j is drawn, which gives the principal the opportunity to cross-train A_1 (without loss of generality) and brings an expected return $\underline{P}\pi + \beta EV_{\min}(j',\theta_1',\theta_2)$ in the case of cross-training. On the other hand, acting in accordance with specializations brings an expected return $\bar{P}\pi + \beta EV_{\max}(j',\theta_1,\theta_2)$. Because cross-training A_i is not optimal now, $\underline{P}\pi + \beta EV^t_{\min}(j',\theta_1',\theta_2) < \bar{P}\pi + \beta EV^t_{\max}(j',\theta_1,\theta_2)$ should hold at t. Now suppose that $\tilde{t} > t$ is the first period in which the signal s_j is drawn again (which again makes cross-training A_1 a possible action). But since the cross-training was not made at t and distribution of the states of the world is time-independent, the dynamic problem of the principal at $\tilde{t}$ is same with the problem at t. Therefore, it directly follows from the inequality above that $\underline{P}\pi + \beta EV^{\tilde{t}}_{\min}(j',\theta_1',\theta_2) < \bar{P}\pi + \beta EV^{\tilde{t}}_{\max}(j',\theta_1,\theta_2)$ holds, as well.

Proof of Theorem 1

We start with the possibility that both cognitive dissonances are reduced to their minimum values, that is, 0. Let's call the first period in which $\theta_i = 0$ for $i = 1, 2$ as $t^{0,0}$. By Lemma 1, $t^{0,0}$ is finite with probability 1. Starting from $t^{0,0}$ onward, the principal's decision is trivial since cognitive dissonances cannot be reduced anymore.[10] Therefore, starting from $t^{0,0}$, the principal acts in accordance with specializations for any given signal in each period. Let's call the expected return from acting in accordance with specializations starting from $t^{0,0}$ onward as $\bar{V}$.

By the definition of $t^{0,0}$, $\exists\theta_i$ such that $\theta_i \neq 0$ at $t^{0,0} - 1$. Since $\theta_i = 0$ at $t^{0,0}$ and $\theta_i \neq 0$ at $t^{0,0} - 1$, we can say that the signal at $t^{0,0} - 1$ was the one in which A_i is not specialized in. So, we can solve the principal's problem backward starting from $\theta_i = 0$, for $i = 1, 2$. If $\theta_i = 0$ is reached for both agents, whatever decision the principal makes, the stochastic processes in the future are same and given the state of the world, the principal should act in accordance with specializations since there is no expected future benefit from attempting cross-training.

Assume that θ_i reaches 0 in n_i^0 successful attempts. Because the principal knows the processes governing θ_is, now when he has value of θ_i that is one successful cross-training away from 0, he can compare the expected return that corresponds to acting in accordance with specializations and attempting cross-training. Given the parameters of the model and processes governing θ_is, the optimal decision rule can be written as a threshold rule that is described in the proposition.

[10] Note that what we actually use here is the fact that they are bounded below by 0, which enables us to use a backward induction argument.

(*Appendix continues*)

Continuing in this fashion, the principal can compare expected returns for each strategy given the state of the world and current values of dissonances at each period t. Therefore, each period decision is determined with the help of a threshold rule presented above. Since we solve the problem in a backward induction fashion, the collection of optimal actions along this path constitute an optimal strategy.

Proof of Corollary 2

Attempting cross-training has expected current costs (e.g., $\pi(\bar{P} - \underline{P})$ if $s_j = s_1$ or s_2) because the principal assigns the signal to the agent who is not specialized in it. On the other hand, it also has expected future benefits if the agent to be cross-trained is open to habituation, that is, $\theta_i(n_i)$ is not a constant function. However, if none of the agents are open to habituation, that is, for $i = 1, 2$, $\theta_i(n_i)$ is a constant dissonance function, then there are no expected future benefits from attempting cross-training. In this case, next period value functions in (1) and (2) are identical for all actions at t. Therefore, the principal's decision is determined only by the current expected return levels. Since in (1), $\bar{P} > \underline{P}$ and in (2), $(\bar{P}\,\bar{P})(\Pi - [\phi_1(\theta_1) + \phi_2(\theta_2)]) > (\underline{P}\,\underline{P})(\Pi - [\phi_1(\theta_1) + \phi_2(\theta_2)])$, the principal does not attempt cross-training if both θ_is are constant. The long run levels of dissonances in this particular case will be the initial values of θ_1 and θ_2.

Proof of Corollary 3

This case is similar to the one discussed in the previous corollary with the only difference that one of the agents is open to habituation. Therefore, the reasoning used in the proofs of Proposition 1 and Corollary 2 is valid again. There is no reason to try to habituate the inflexible agent and the decision on the other agent depends on other parameters. The long run level of dissonance for the inflexible agent is its initial value whereas for the flexible agent, the decision rule defined in the main theorem combined with parameter values determines the long run value.

Proof of Proposition 1

The frequency of the composite state affects principal's decision because it directly affects the expected benefit from attempting cross-training. F_j terms are hidden in the expectation operator in value function, (1) and (2). In particular, an increase in the probability of a conflict means a higher marginal expected future return for the cross-training. Accordingly, if $\tilde{F}_3 > F_3$, then $\bar{\tilde{\theta}} < \bar{\theta}_i$ because $E\widetilde{V}(j,\bar{\tilde{\theta}}_1,\bar{\tilde{\theta}}_2) > E\widetilde{V}(j,\bar{\theta}_1,\bar{\theta}_2)$ for cross-training. Hence, an increase in the probability of s_3 (i.e., F_3) implies a greater tendency for attempting cross-training.

Proof of Proposition 2

We initially assume that for all $i, \dfrac{\Delta\theta_i}{\Delta n_i} < 0$ at each point, where n_i is the number of successful cross-training attempts for agent A_i and θ_i is a function defined on natural numbers. Now, assume for all $i, \dfrac{\Delta\tilde{\theta}_i}{\Delta n_i} < 0$, $\tilde{\theta}_i(n_i) \leq \theta_i(n_i)$ and, $\dfrac{\Delta\tilde{\theta}_i}{\Delta n_i} < \dfrac{\Delta\theta_i}{\Delta n_i}$ at each point n_i. An increase in habituation rate means that (i) higher (potential) returns in the composite state (s_3) will be obtained sooner, that is, the cross-training investment starts to pay off sooner, and (ii) the opportunity cost of reaching any given level of dissonance is now lower, both of which imply a greater tendency for attempting cross-training. The result follows.

(*Appendix continues*)

Proof of Proposition 3

The logic behind this result is very intuitive. An increase in patience or forward-lookingness (i.e., an increase in β) means a higher valuation of future benefits, which can be obtained by cross-training. To show this, consider the problem of the principal at any point in time t for any level of θ_i, under signal s_j which makes it possible to cross-train agent A_1 (without loss of generality). Assume that $\tilde{\beta} > \beta$. Denote the expected future returns if cross-training is made by $EV\,(j,\,\theta'_1,\,\theta_2)$ and the expected future returns if specializations are followed by $EV\,(j,\,\theta_1,\,\theta_2)$. As we know, if

$$\beta\big(EV\big(j,\theta_1^{'},\theta_2\big)-EV(j,\theta_1,\theta_2)\big)>(\overline{P}-\underline{P})\,\pi$$

then it is optimal to attempt cross-training. But,

$$\begin{aligned}&\tilde{\beta}\big(EV\big(j,\theta_1^{'},\theta_2\big)-EV(j,\theta_1,\theta_2)\big)\\&\quad>\beta\big(EV\big(j,\theta_1^{'},\theta_2\big)-EV(j,\theta_1,\theta_2)\big)>(\overline{P}-\underline{P})\,\pi\end{aligned}$$

because $\tilde{\beta} > \beta$ and $EV\,(j,\,\theta'_1,\,\theta_2) - EV\,(j,\,\theta_1,\,\theta_2) > 0$. Because this is valid for any t, s_j, and θ_i, it is also valid for long run values of θ_i, which implies $\tilde{\bar{\theta}} < \bar{\theta}_i$ if $\tilde{\beta} > \beta$.

Received September 18, 2012
Revision received September 5, 2013
Accepted September 8, 2013 ■